J. Mol. Biol. (1979) **133**, 359–383

Construction, Propagation and Expression of Simian Virus 40 Recombinant Genomes Containing the *Escherichia coli* Gene for Thymidine Kinase and a *Saccharomyces cerevisae* Gene for Tyrosine Transfer RNA

STEPHEN P. GOFF† AND PAUL BERG

Department of Biochemistry
Stanford University School of Medicine
Stanford, Calif. 94305, U.S.A.

(Received 28 December 1978)

Recombinant simian virus 40 (SV40) virus genomes have been constructed *in vitro* by joining SVGT1, the segment of SV40 DNA between map co-ordinates 0·15 and 0·73 (clockwise), to either the *Escherichia coli* gene for thymidine kinase (*Ecotdk*), or one of the *Saccharomyces cerevisae* genes for $tRNA^{Tyr}$ (*ScetyrG*); the resulting recombinants were propagated in CV-1 monkey cells at 41°C using *ts*A58 as a helper. The *Ecotdk* gene was obtained first as an approximately 20 kb‡ DNA segment in a defective transducing phage, *ϕ80dtdk5*, then as overlapping 1·85 and 2·35 kb segments in pMB9-*Ecotdk*, prior to insertion into SVGT1. The *ScetyrG* segment introduced into SVGT1 was a 1·25 kb DNA fragment contained in λgtl-*ScetyrG*, a recombinant that had been cloned previously by Olson *et al.* (1979).

After infection of CV-1 monkey cells with the SVGT1-*Ecotdk* or SVGT1-*ScetyrG* hybrid viruses, RNA complementary to the exogenous DNA was produced. With SVGT1-*Ecotdk* the RNA homologous to the *Ecotdk* segment was heterogeneous, ranging in size from 1 to > 8 kb in length. At least 30 to 40% of this RNA was covalently joined to SV40-specific RNA. No *E. coli* thymidine kinase enzyme activity could be detected in the infected cells. By contrast, infection with SVTG1-*ScetyrG* resulted in the formation of a transfer RNA-sized RNA, complementary to the $tRNA^{Tyr}$ coding sequence of the *ScetyrG* segment, as well as a population of heterogenous large RNAs. Since the formation of the tRNA-sized RNA occurred after infection with recombinants having the *ScetyrG* segments in either of the two alternative orientations in SVGT1, it is possible that transcription of the *tyrG* sequence is initiated within the cloned segment.

1. Introduction

Transducing bacteriophages carrying specific regions of the bacterial chromosome have contributed significantly to the analysis of bacterial genome organization, expression and regulation. Phage-host recombinants provided novel cellular and phage genotypes, starting materials for detailed structure and base sequence analyses, specific hybridization probes for monitoring gene expression, and abundant specific gene products, i.e. RNAs and proteins, for investigations of the mechanisms of

† Present address: Massachusetts Institute of Technology, Cancer Center, Cambridge, Mass., U.S.A.

‡ Abbreviation used: kb, kilobase (10^3 bases).

0022 2836/79/270359-25 $02.00/0

transcription and translation. Not surprisingly, therefore, the growing interest and studies in eukaryote genome organization, expression and regulation prompted a search for comparable transducing capabilities for animal cells.

Several years ago we (Goff & Berg, 1976) and others (Ganem *et al.*, 1976; Nussbaum *et al.*, 1976; Davoli *et al.*, 1976; Hamer, 1977) explored the feasibility of using simian virus 40 (SV40) as a transducing virus for mammalian cells. To increase the probability of obtaining viral recombinants carrying specific genetic elements, we constructed the recombinant DNA molecules *in vitro*. This involves joining suitably modified segments of an exogenous DNA to subgenomic portions of SV40 DNA and propagating the recombinant viruses in cultured monkey cells (Goff & Berg, 1976). An obligatory feature of this experimental design is that the SV40 DNA vector must contain the origin of SV40 DNA replication (at map position 0·67). Moreover, because the recombinant genomes lack the genetic function(s) coded by the viral DNA segment that is excised, they must be propagated by complementation with helper virus genomes that can supply the missing gene's product(s). In our experimental design the vector DNA retains at least one functioning gene and, therefore, can complement a defective gene in the helper virus. For example, replacement of the late region of SV40 by an exogenous DNA segment permits the recombinant to complement and be complemented by a *ts* early mutant (*ts*A58) at high temperature, a condition in which neither the recombinant nor *ts*A58 can grow alone (Mertz & Berg, 1974; Goff & Berg, 1976). Currently, our protocol requires that the size of the recombinant DNA molecule be within the packaging limit of SV40 DNA, thereby restricting the size of the exogenous DNA segment that can be propagated with SV40 as the vector.

A principal goal of using SV40 as a transducing virus is to obtain expression of the exogenous DNA segments in cells infected by the recombinants. In principle that expression could be regulated by transcriptional promoters, terminators, processing sites, translational signals, etc., on the SV40 DNA, on the exogenous DNA segment or both. To explore these possibilities we have constructed two recombinant viral genomes in which most of the late region of SV40 DNA is replaced with a gene from *Escherichia coli* specifying the enzyme thymidine kinase (*Ecotdk*), or a gene from *Saccharomyces cerevisae* coding for $tRNA^{Tyr}$ (*ScetyrG*); these hybrid viruses were propagated in CV-1 monkey cell cultures at 41°C with *ts*A58 as the helper. Following infection of CV-1 cells with SVGT1-*Ecotdk*, high molecular weight heterogenous RNA complementary to the *tdk* segment is produced; however, we have not detected either specific cytoplasmic messenger RNAs containing the sequence of *Ecotdk* or thymidine kinase activity in the infected cells. By contrast, cells infected with SVGT1-*ScetyrG* produce a discrete RNA species whose electrophoretic mobility in agarose gel is very nearly the same as tRNAs and contain the nucleotide sequence of the *ScetyrG* gene.

2. Materials and Methods

(a) *Animal cells and viruses*

The growth of CV-1 and CV-1P monkey cells has been described by Mertz & Berg (1974) and the African green monkey cell line lacking thymidine kinase activity, T22 TK^-, was obtained from C. Croce (Wistar Institute) and propagated in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum with 30 μg BrdUrd/ml. The origin and propagation of our wild-type strain of SV40 (SVS) and *ts*A58 virus has been reported

(Mertz & Berg, 1974); the procedures for transfection, detection, purification and propagation of the viruses containing recombinant genomes were described previously (Goff & Berg, 1976).

(b) *Bacterial and phage growth and induction of lysogens*

Wild-type $\phi 80$ phage was obtained from A. D. Kaiser (Stanford University), *E. coli* strain KY895 (W3110 *tdk*$^-$) from B. Bachmann (New England Culture Center) and both *E. coli* W3110 and HB101, containing the pMB9 plasmid (Bolivar *et al.*, 1977), from R. W. Davis (Stanford University). λgtl clones containing the *S. cerevisiae* genes *tyrG* and *tyrE*, as well as recombinants of these genes with pMB9 (pMB9-*ScetyrG* and pMB9-*ScetyrE*, henceforth abbreviated to p*tyrG* and p*tyrE*), were those described by Olson *et al.* (1979). Transfection of *E. coli* strain KY895 was performed by the $CaCl_2$ method as described by Wensink *et al.* (1974).

Bacteria were usually grown in L-broth (5 g NaCl, 5 g yeast extract, 10 g tryptone/l, pH 7·5); to select thymidine kinase-positive (TK$^+$) cells the medium also contained 25 μg fluorodeoxyuridine/ml, 25 μg uridine/ml, and 50 μg thymidine/ml. Tetracycline-resistant (tet^R) cells were selected for growth on 15 μg tetracycline/ml. To induce $\phi 80$ lysogens, cells were grown to an A_{600} of 1 unit, centrifuged, resuspended in an equal volume of buffer (10 mM-Tris, pH 7·5, 10 mM-$MgCl_2$) and exposed to u.v. light for 10 s in a large flat pan. After centrifugation the cells were resuspended in fresh medium and grown at 37°C for 5 h with vigorous shaking. The lysate was treated with $CHCl_3$, clarified by centrifugation (10 min at 10,000 revs/min) and the phage in the supernatant were purified further by 2 successive centrifugations in CsCl density-step gradients and then to equilibrium in a CsCl density gradient.

(c) *Preparation of phage, plasmid, and viral DNA*

DNA was prepared from purified phage by the formamide lysis method (Cameron, 1977). To obtain the *Ecotdk* gene for introduction into pMB9, the *ϕ80dtdk5* DNA was sheared to smaller fragments: 40 μg of a mixture of $\phi 80$ and *ϕ80dtdk5* DNA, in a total vol. of 25 ml of TENN buffer (10 mM-Tris, pH 7·5, 1 mM-EDTA, 1 M-NaCl) was stirred in a Virtis homogenizer at 15,000 revs/min for 1·5 h at 4°C, and the fragmented DNA was concentrated by precipitation with alcohol. These conditions produce DNA with an average size of about 2 kb† as judged by agarose gel electrophoresis.

Supercoiled plasmid DNA was prepared by lysozyme/EDTA/Triton X100 lysis (Katz *et al.*, 1973) and SV40 DNA by differential extraction from infected CV-1 cells (Hirt, 1967). Both DNAs were purified by equilibrium centrifugation in CsCl containing ethidium bromide as described by Radloff *et al.* (1967).

(d) *Preparation of ^{32}P-labeled DNA*

DNA was nick-translated according to Rigby *et al.* (1977) using 0·5 to 2 μg of SV40, pTK1, p*tyrG* or p*tyrE* DNAs and [α-^{32}P]dGTP (New England Nuclear) as the labeled substrate. The specific activities of the DNAs were between 10^7 to 10^8 cts/min per μg.

(e) *Hybridizations*

To locate or identify SV40 recombinant plaques on CV-1P cells, the infected cell monolayers were transferred to nitrocellulose discs, hybridized with the appropriate ^{32}P-labeled nick-translated DNAs and autoradiographed (Villarreal & Berg, 1977). To confirm these initial identifications, supernatants obtained by Hirt's (1967) procedure from cells infected with the initial virus stock or plaque-purified virus, were spotted (1 μl of each supernatant) directly onto a square of dry Schleicher and Schuell nitrocellulose and hybridized to appropriate ^{32}P-labeled DNA as described by Villarreal & Berg (1977); this procedure simplified and hastened the screening of hundreds of putative recombinant virus clones.

^{3}H-labeled RNA purified from CV-1 cells infected with the different recombinant viruses was annealed to the appropriate DNA loaded onto nitrocellulose discs. SV40 DNA was cleaved with *Taq*I endonuclease and pTK1 and p*tyrG* DNAs were cleaved with

† See footnote to p. 359.

*Hin*dIII endonuclease, prior to alkali denaturation and loading onto filters (5 μg/filter). After annealing, the nitrocellulose discs were washed and counted as previously described (Goff & Berg, 1976). For rehybridization the bound RNA was eluted from the filters by heating to 100°C for 5 min in 1 ml of TE buffer (10 mM-Tris, pH 7·5, 1 mM-EDTA), and then precipitated with ethanol.

The size of RNA molecules made after infection with recombinant viruses was analyzed by hybridization after the glyoxal-denatured RNA had been electrophoresed in 1·5% or 2·5% agarose gels (McMaster & Carmichael, 1977). The gel was treated with 50 mM-NaOH for 50 min at 20°C, neutralized by washing twice with 200 mM-$NaPO_4$ (pH 6·8) and twice with 20 mM of the phosphate buffer; then, the RNA was transferred from the gels onto diazobenzyloxymethyl paper (Alwine *et al.*, 1977) to which the RNA binds covalently, thereby maintaining the pattern produced by electrophoresis. The paper was washed and annealed with the appropriate ^{32}P-labeled DNA probe as described above.

(f) *Electrophoresis of DNA*

Restriction endonuclease digestion products were analyzed by electrophoresis in agarose gels as previously described (Goff & Berg, 1976) using known DNA fragments as molecular length standards. To prepare specific DNA fragments, up to 100 μg of DNA were loaded in a horizontal slab gel of 0·7% agarose (0·6 cm × 10 cm × 21 cm) in TBE buffer (Peacock & Dingman, 1968) and electrophoresed overnight at 1 V/cm. DNA was recovered from appropriate agarose slices by electrophoretic elution in 20-fold diluted TBE buffer.

(g) *Electron microscopy*

Heteroduplexes for electron microscopy were prepared according to Davis *et al.* (1971), except that DNAs containing poly(dA) and poly(dT) sequences were mounted from solutions containing 70% formamide onto a hypophase containing 40% formamide to prevent annealing of the homopolymer sequences.

(h) *Assay for thymidine kinase activity*

A 10-ml bacterial culture was grown overnight in L-broth (containing 15 μg tetracycline/ml for plasmid-carrying strains) with vigorous aeration at 37°C, centrifuged, washed twice with 25 mM-Tris buffer (pH 7·4) containing 120 mM-NaCl, 5 mM-KCl, 1 mM-$CaCl_2$, 1 mM-$MgCl_2$, and 1 mM-Na_2HPO_2 and resuspended in 0·3 ml of 10 mM-Tris buffer (pH 8·0) with 150 mM-KCl and 3 mM-β-mercaptoethanol; the cells were sonicated 4 times for 45-s intervals while cold. The extract was centrifuged at 19,000 revs/min for 30 min at 4°C and the supernatant was heated to 70°C for 5 min with occasional shaking. After cooling, the extract was centrifuged at 10,000 revs/min for 15 min at 4°C and stored frozen at −70°C. Similar extracts of CV-1 monkey cells were prepared 48 h after infection.

The assay mixture (20 μl) contained 100 mM-Tris buffer (pH 8·0), 10 mM-ATP, 12 mM-$MgCl_2$, 12 mM-3-phosphoglycerate, 110 μM-[^{3}H]dThd (spec. act. 1·8 Ci/mmol). After adding appropriate dilutions of the extract (30 μl) and incubation for various times at 37°C, the reaction was stopped by adding 300 μl of cold ammonium acetate buffer (1 mM, pH 8·0) containing 10 mM unlabeled dThd and the mixture was poured slowly through DEAE-cellulose discs (DE-81; Whatman). The discs were washed successively with 15 ml of 1 mM-ammonium acetate (pH 8·0), and 5 ml of ethanol, then dried and counted in a toluene-based liquid scintillant. The efficiency of counting [^{3}H]dTMP is approximately 5% under these conditions.

(i) *Assay for thymidine kinase in polyacrylamide gel after electrophoresis*

Thymidine kinase was also assayed following electrophoresis of crude cell extracts in polyacrylamide gels as described by Tischfield *et al.* (1973). A vertical slab gel (30 ml), containing 7% acrylamide, 0·2% *N,N'*-methylene bisacrylamide, 400 mM-Tris buffer (pH 8·5), 2 mM-$MgCl_2$, 0·06% TEMED, and 5 μg riboflavin/ml, was photopolymerized for 2 h. A spacer gel containing 4% acrylamide, 0·1% *N,N'*-methylene bisacrylamide, 100 mM-Tris buffer (pH 8·0), 1 mM-$MgCl_2$, 0·1% TEMED, and 4 μg riboflavin/ml was

poured on top of the separating gel and photopolymerized for 1 h. The tray buffers contained 250 mM-Tris-OH, 1·9 M-glycine; the upper tray buffer also contained 2·5 mM-ATP and 0·1 mM-dCDP. The extract (40 μl) contained 25 mM-Tris buffer (pH 8·5), 2·5 mM-$MgCl_2$, 6 mM-ATP, 2·5 mM-dCDP, and 5% glycerol; electrophoresis was for 4 h at 4°C at a constant current of 35 mA. The gel was soaked in 10 ml of the assay mixture containing 100 mM-Tris buffer (pH 7·0), 10 mM-ATP, 10 mM-$MgCl_2$, and 50 μM-[^{14}C]dThd (50 mCi/mmol) for 2 h at 37°C, fixed for 12 h at 4°C in 250 ml of 100 mM-Tris buffer (pH 7·0), containing 100 mM-lanthanum chloride, and washed in running water for 24 h. Finally, the gel was dried and autoradiographed.

The construction, propagation and handling of SVGT1-*Ecotdk* and SVGT1-*ScetyrG* recombinants were performed in a P3 facility as recommended in the National Institutes of Health guidelines.

3. Results

(a) *Rationale*

Attempts to clone a particular gene or segment of DNA from a source in which it is represented only rarely have generally resorted to the "shotgun" approach: random cleavage of the starting DNA using mechanical shear or restriction endonucleases, ligation of these fragments to a suitable vector DNA, followed by selection or screening for recombinants that express the particular biological function or contain the particular nucleotide sequence. But the efficiency of recovering recombinants with SV40 DNA is low and the selection-screening procedures for specific genes using animal cell cultures are limited, tedious, expensive, or non-existent. Consequently, it is preferable, first, to trim and purify the DNA segment to the desired state by one or more clonings in bacteria and then to transfer the cloned DNA segment from the bacterial vector to an appropriate SV40 vector. This two-step procedure was used to construct recombinant molecules with SV40 DNA and DNA segments containing *E. coli*'s gene for thymidine kinase (*Ecotdk*) or a *S. cerevisae* gene for $tRNA^{Tyr}$ (*ScetyrG*). The SV40 vector was SVGT1, a 3·05 kb segment of SV40 DNA containing the origin of DNA replication and the entire early region (the segment defined by map co-ordinates 0·73 counter-clockwise to 0·15) (Goff & Berg, 1976). SVGT1 was produced by successive cleavages with *Mno*I and *Bam*HI endonucleases and purified by electrophoresis on an agarose gel; its 3′ ends were modified by the addition of poly(dA) (about 50 to 100 nucleotides long) as previously described (Goff & Berg, 1976).

(b) *Construction of SVGT1*-Ecotdk *recombinants* *(see Fig. 1 for an overall summary)*

(i) *Preparation of ϕ80*dtdk *phages*

E. coli tdk maps close to the attachment site for ϕ80 (Bachmann *et al.*, 1976). Consequently, ϕ80*dtdk* transducing phages could be obtained after ultraviolet light induction of a ϕ80 lysogen of *E. coli* W3110 (*tdk*$^+$). These transducing phages were detected by infecting *E. coli* KY895 (*tdk*$^-$) and plating the infected cells in the presence of FdUrd plus uridine and dThd. Under these conditions less than 10^{-6} of uninfected cells form colonies (because they cannot phosphorylate thymidine (TK^-)), but approximately 10^{-5} of infected cells become TK^+ and produce colonies on this selective medium. Eight of these transductants were grown, induced with ultraviolet light and the phage lysates tested for *tdk* transducing activity as described above.

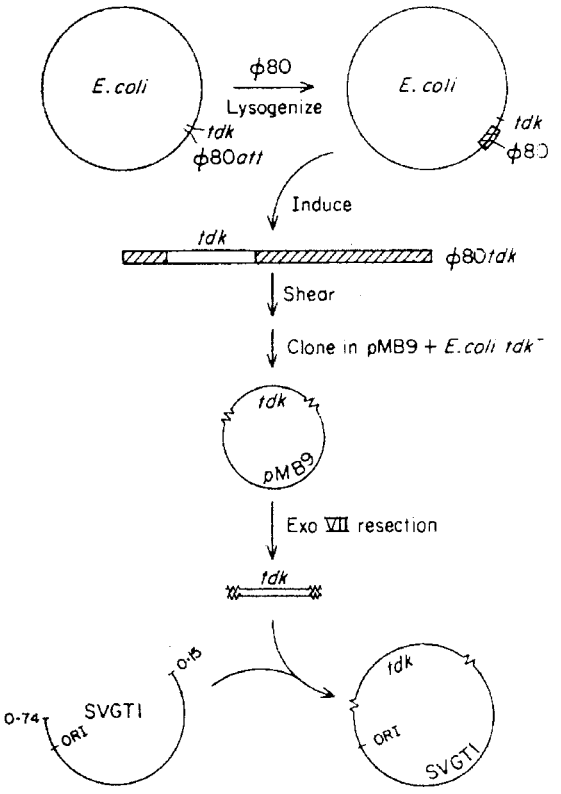


FIG. 1. Scheme for cloning a small DNA fragment containing the *E. coli tdk* gene in SVGT1.
A ϕ80 lysogen of *E. coli* was constructed, and u.v. irradiation was used to induce a lysate which could transduce *E. coli* TK$^-$ to TK$^+$. u.v. induction of one of these TK$^+$ transductants produced ϕ80*dtdk5*. The ϕ80*dtdk5* phage DNA was sheared and the resulting fragments were cloned in *E. coli* using pMB9 DNA as the vector. pMB9-*Ecotdk* recombinant plasmids were detected by transformation of *E. coli* TK$^-$ to TK$^+$. The *tdk* segment was excised (Goff & Berg, 1978), joined to SVGT1 and the resulting SVGT1-*Ecotdk* recombinants were propagated in CV-1 cells.

Seven of these phage lysates transduced KY895 from TK$^-$ to TK$^+$ at frequencies between 10^{-1} to 10^{-2} after infection at a multiplicity of ten.

ϕ80*dtdk5*, one of the seven transductants, was obtained following ultraviolet light induction of a 200-ml culture and centrifugation in a CsCl density gradient: the transducing phage was present in the denser of the two bands. Judging from the structure of heteroduplexes between wild-type ϕ80 and the transducing phage's DNA (not shown), about $13{\cdot}9 \pm 1{\cdot}0$ kb of the left half of ϕ80 DNA is replaced by $19{\cdot}5 \pm 1{\cdot}2$ kb of *E. coli* DNA. Thus, *Ecotdk* maps less than 20 kb from *E. coli*'s ϕ80 attachment site.

(ii) *Preparation of pMB9*-Ecotdk

ϕ80*dtdk5* DNA was sheared to an average size of 2 kb and poly(dT) "tails" were added to the fragments; these were annealed to *Eco*RI endonuclease-cleaved pMB9 DNA whose termini contained poly(dA) tails. *E. coli* KY895 was then transformed with the annealed DNA using the $CaCl_2$ method (Wensink *et al.*, 1974) and tetracycline-resistant (tet^R) clones were selected. Between $0{\cdot}5 \times 10^3$ to 2×10^3 tet^R colonies per μg of annealed DNA were found in four separate experiments. Two of the

320 tet^R colonies tested were TK^+ as judged by their ability to grow on FdUrd-containing media. One contained an insert of 1·85 kb DNA (pTK1) and the other's insert was 2·35 kb DNA (pTK2).

Both inserts have three *Hinc*II endonuclease cleavage sites (see Fig. 2) and this allows the maps of pTK1 and pTK2 to be aligned with each other. pTK1 contains two *Eco*RI restriction sites spaced 0·2 kb apart near one end of the cloned segment while pTK2 contains only one of them; the second *Eco*RI restriction site must lie outside the sheared fragment inserted in pTK2. After aligning all of the shared restriction sites, it is clear that the two independently cloned DNA segments share a region of about 1·5 kb, a length only slightly larger than is needed to code for the thymidine kinase subunit of 40 kilodaltons (Okazaki & Kornberg, 1964*a*,*b*).

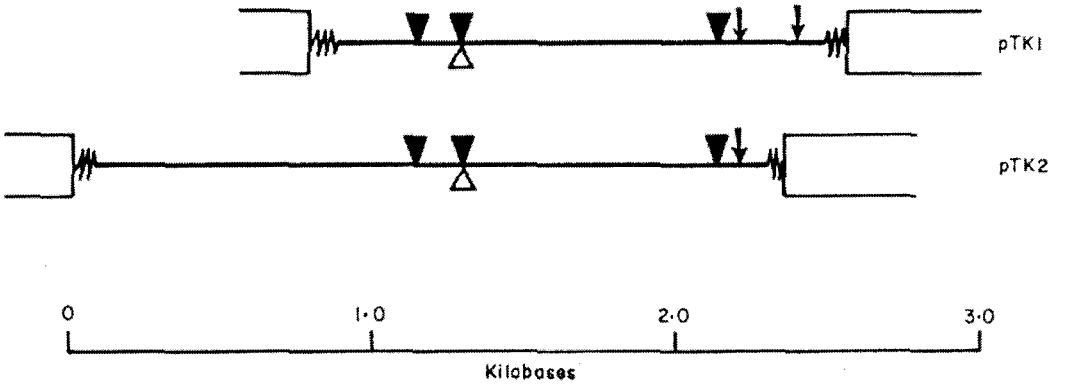
DNA heteroduplexes between linear pTK1 and pTK2 DNA molecules (the two recombinant DNAs are cleaved once by *Bam*H1 endonuclease in the pMB9 portion) showed that their respective inserts have the same orientation (Fig. 2). The plasmid and *E. coli* DNA segments are almost completely base-paired in the heteroduplexes, the only non-homology between them being the small loop of DNA representing the left end of the pTK2 segment that is missing from the pTK1 insert. From the length of this loop we infer that the pTK2 cloned DNA segment contains 0·7 kb more DNA to the left of the pTK1 insert (see Fig. 2); the unshared 0·2 kb at the right end of pTK1 is too small to detect by heteroduplex analysis under these conditions.

(iii) *Expression of the cloned* tdk *DNA segments in* E. coli

Since the frequency of TK^+ clones amongst the tet^R transfectants was low, it was important to establish that the TK^+ phenotype did not result from reversion of KY895 and that the cloned DNA segments did indeed carry a functional *Ecotdk* gene. pTK1 and pTK2 plasmid DNAs were isolated and used to transform KY895 cells as mentioned above. With pMB9 plasmid DNA, 0/15 tet^R clones were TK^+; but with pTK1 and pTK2 DNAs, 18/18 and 10/10 tet^R colonies, respectively, became TK^+.

Cells harboring the pTK1 plasmid also produce thymidine kinase activity (Fig. 3). KY895 cell extracts have only 2% as much thymidine kinase activity as do *E. coli* W3110, but KY895 cells carrying pTK1 have 5·7 times more thymidine kinase than the wild-type cells.

To determine if the *tdk* gene in pTK1 contains all the information needed for thymidine kinase synthesis, or if the expression of the *tdk* gene is governed by its placement within the pMB9 vector, the cloned segment in pTK1 was excised with its poly(dT) tails intact (Goff & Berg, 1978) and annealed to linear pMB9 DNA containing poly(dA) tails at its *Eco*RI endonuclease-generated ends. After transformation of *E. coli* KY895, TK^+ clones were selected and their plasmid DNAs isolated. Of six recombinant plasmids examined by cleavage with *Sal*I endonuclease (one *Sal*I restriction site occurs in the plasmid and one in the cloned *tdk* DNA) one contained the *Ecotdk* DNA segment with the same orientation as pTK1 and five had the opposite orientation. These assignments were confirmed by digestion with both *Hin*dIII and *Eco*RI endonucleases. This experiment shows that the *tdk* gene can be expressed irrespective of its orientation in the plasmid vector; all of the signals needed to express thymidine kinase activity are probably contained within the cloned DNA segment.



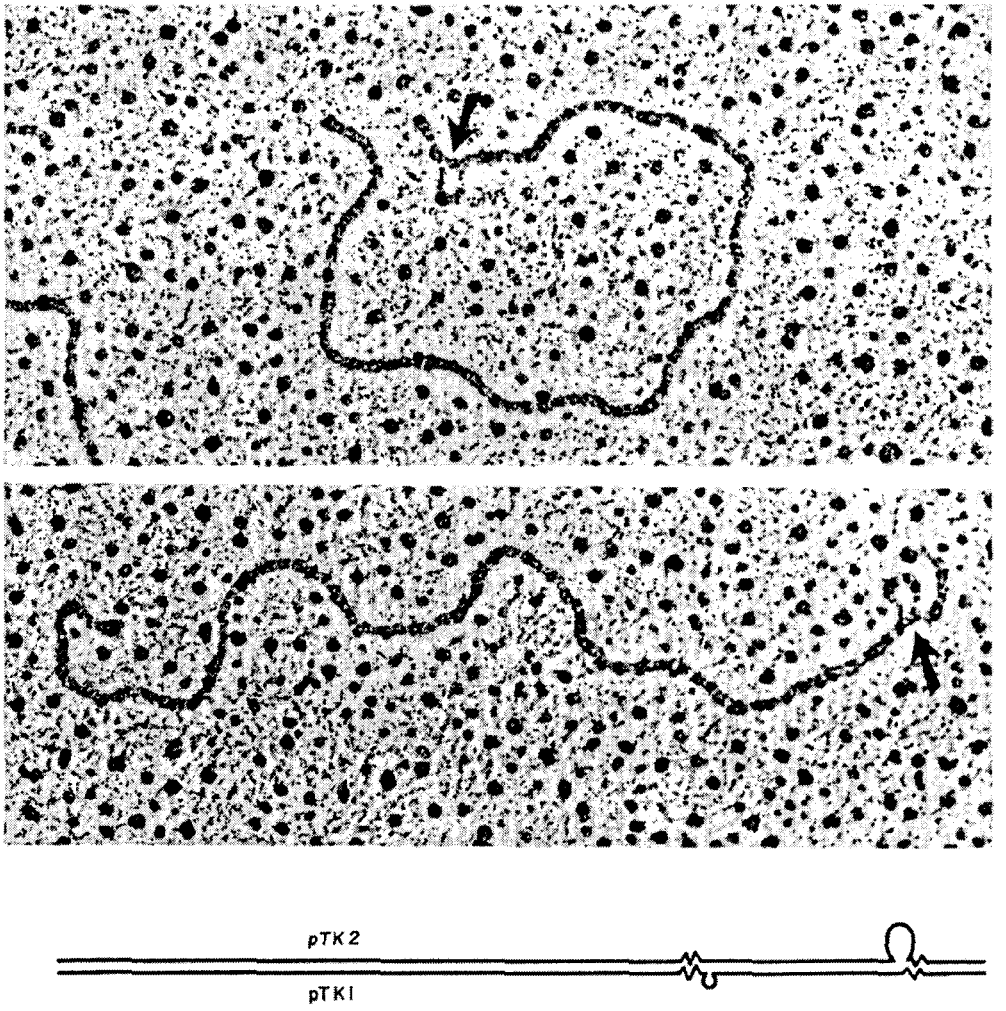


FIG. 2. *Eco*RI (↓), *Hinc*II (▼) and *Sal*I (△) restriction endonuclease cleavage maps of the *Ecotdk* DNA segments in pTK1 and pTK2 plasmids (top) and electron micrographs of heteroduplexes between pTK1 and pTK2 DNAs (bottom).

The diagram below the electron micrographs of the heteroduplexes formed from linear pTK1 and pTK2 DNAs symbolizes the shared regions as straight parallel lines, the non-homologous regions as loops; the jagged lines represent the dA:dT join sequences.

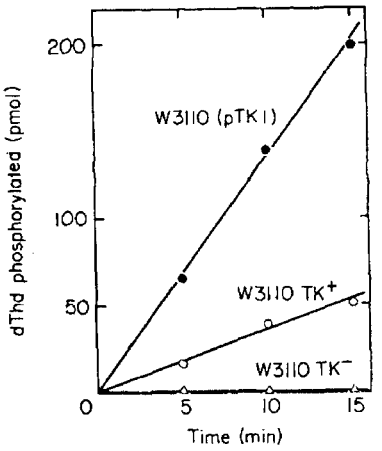


FIG. 3. Thymidine kinase activity in extracts of *E. coli* W3110 (TK$^+$) (○), KY895 (TK$^-$) (△) and KY895 (pTK1) (●).

Extracts of the bacterial cells were prepared and assayed for thymidine kinase activity as described in Materials and Methods.

(iv) *Construction and isolation of SVGT1*-Ecotdk

During transfer from *E. coli*'s genome to the pTK1 plasmid the *Ecotdk* segment was purified 2000-fold; it was also trimmed to a size that can be accommodated within the SVGT1 vector. To construct SVGT1-*Ecotdk* the *Ecotdk* segment was excised from pTK1 with its poly(dA:dT) flanking sequences intact (Goff & Berg, 1978) and after brief digestion with λ 5′-exonuclease to regenerate poly(dT) tails it was annealed to SVGT1 bearing poly(dA) tails. The resulting mixture, with added *ts*A58 DNA, was used to transfect CV-1P cells and, after 12 days, the resultant plaques were screened by *in situ* hybridization for *Ecotdk* nucleotide sequences (Villarreal & Berg, 1977). Approximately half of the 10^4 plaques formed per μg of annealed DNA were *Ecotdk* positive. CV-1 cultures were infected with virus from 60 randomly picked plaques and the extracted viral DNA (Hirt, 1967) was screened for *Ecotdk* sequences by hybridizing DNA spots on nitrocellulose to ^{32}P-labeled pTK1 DNA; nearly half the plaques contained recombinant genomes. A total of 24 of the original plaques containing SVGT1-*Ecotdk* recombinants were plaque-purified once on CV-1P cells using added *ts*A58 as helper virus; two purified plaques from each of the original 24 were used to prepare virus stocks and viral DNA. A total of 35 of these 48 DNAs hybridized strongly with ^{32}P-labeled pTK1 DNA indicating that they contain sequences homologous to the *Ecotdk* DNA segment in pTK1.

(v) *The structure of the cloned SVGT1*-Ecotdk *DNAs*

The mixture of recombinant and *ts*A58 DNA in each of the 35 DNA preparations mentioned above was digested with *Eco*RI and *Hin*dIII endonucleases and the resulting fragments were electrophoresed on a 1·4% agarose gel (Fig. 4). Ten of the DNAs yielded the expected pattern of DNA fragments: in addition to the fragments expected from wild-type SV40 DNA (those contributed by the *ts*A58 helper DNA present in the preparations) there were two new bands. The sizes of these bands are consistent with those predicted for the fragments that contain both SV40 and *tdk* sequences and span the poly(dA:dT) join (Fig. 4). The two new fragments are either approximately 1·85 kb and 1·1 kb, or 2·35 kb and 0·6 kb in length. These fragment sizes are those expected for the two possible orientations of the *tdk* DNA segment

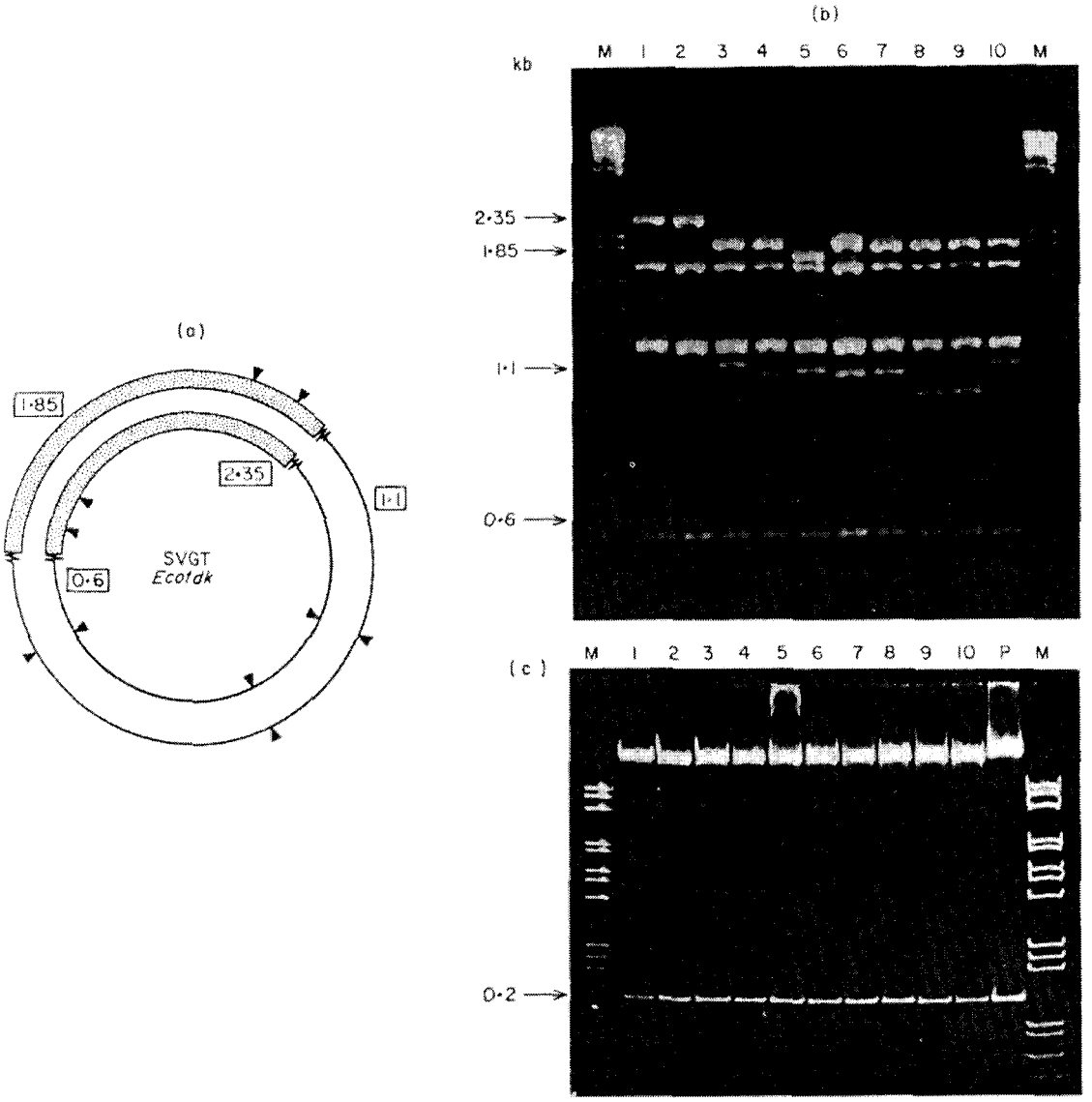


FIG. 4. Electrophoretic analysis of *Eco*RI and *Hin*dIII endonuclease digests of SVGT1-*Ecotdk* recombinant DNAs.

(a) *Eco*RI and *Hin*dIII restriction map of SVGT1-*Ecotdk* DNAs having the two possible orientations of the *Ecotdk* segment.

(b) Electrophoretic separation of the larger DNA fragments on agarose gel. The outermost tracks contain molecular length standards produced by digestion of phage λ DNA with *Eco*RI and *Hin*dIII endonucleases. Tracks 1 and 2 show DNAs yielding the 2·35 and 0·6 kb fragments indicative of the orientation in the inner ring of (a); tracks 3 to 10 are DNAs that generate the 1·85 and 1·1 kb fragments, indicative of the orientation shown in the outer ring. Tracks 1 to 10 contain, respectively, recombinant clones 50B, 50A, 47B, 44A, 28A, 27B, 27A, 22B, 22A, and 6A.

(c) Electrophoretic separation of the small fragments produced by *Eco*RI endonuclease digestion. The outermost tracks are length standards produced by *Hae*III endonuclease digestion of phage PM2 DNA. Tracks 1 to 10 contain the same recombinant DNAs shown in (b) and track P contains pTK1 DNA cleaved with *Eco*RI endonuclease. The 0·2 kb fragments produced in all tracks have identical mobilities.

relative to the SV40 vector. Thus, eight of the hybrids (of which only five are independent) have the *tdk* DNA segment inserted in one orientation, while two (both from one original plaque) have the same segment in the reverse orientation (Fig. 4). The sizes of the new fragments produced by the *Eco*RI plus *Hin*dIII endonuclease digestions are different in each clone. This variation probably results from differences in the length of the poly(dA:dT) join sequences; indeed, the variation is restricted to the fragments containing the poly(dA:dT) joins, while the sizes of the internal fragments are constant.

The ten recombinants referred to above also produced the expected 0·2 kb fragment from within the *Ecotdk* segment after *Eco*RI endonuclease cleavage (Fig. 4(c)). As expected, the 0·2 kb fragment is indistinguishable from the small fragment produced by *Eco*RI endonuclease digestion of the plasmid pTK1; moreover, the same fragment is produced from SVGT1-*Ecotdk* recombinants that have the *Ecotdk* segment in each of the two orientations.

Based on the amount of DNA removed from the SV40 vector (2·19 kb) and the size of the *Ecotdk* DNA segment inserted in its place (1·85 kb), the SVGT1-*Ecotdk* recombinants should be about 0·93 SV40 fractional length. Electrophoresis of mixtures of recombinant and helper supercoiled DNAs in 1·5% agarose gel revealed two bands, one conforming to the size of wild-type SV40 DNA and a smaller species about 0·90 to 0·95 SV40 fractional length. The amount of the smaller species ranged between 20 to 50% of the sum of the two DNAs, indicating that the recombinant and helper genomes were nearly equally proficient in their multiplication.

DNA from the remaining 25 recombinant virus isolates yielded heterogeneous and unexpected fragments, many in less than molar amounts, when cleaved by *Eco*RI and *Hin*dIII endonucleases; these have not been characterized further. Quite possibly these genomes have deletions or duplications that arose during their propagation, a phenomenon known to occur with wild-type and deletion mutant viruses (Mertz & Berg, 1974). Alternatively, these "scrambled" genomes could have been formed in the original recombination event and were complemented by the helper virus during the plaque purifications and subsequent growth.

(vi) *Transcription and translation of the cloned* Ecotdk *DNA segment in monkey cells*

Is the *tdk* gene expressed following infection of monkey cells with the SVGT1-*Ecotdk*? Transcription of the *Ecotdk* DNA segment was examined first. CV-1 cell monolayers were infected (3 to 10 plaque-forming units/cell) with SV40 (SVS), or with *ts*A58 and either of two SVGT1-*Ecotdk* isolates (the *Ecotdk* segments in each were oriented in opposite directions). After 48 hours, [^{3}H]uridine (30 μCi/ml; 25 Ci/mmol) was added and four hours later total RNA was isolated and hybridized to nitrocellulose discs containing SV40 or pTK1 DNA (Fig. 5). As expected, RNA obtained from each of the infected cell cultures contained SV40-specific nucleotide sequences. The RNA extracted from SV40 or mock-infected cells did not hybridize to pTK1 DNA, but RNA obtained from cells infected with either of the two SVGT1-*Ecotdk* recombinants did hybridize to pTK1 DNA. Since more than 95% of the ^{3}H label eluted from the filters became acid-soluble after a brief incubation with alkali, we conclude that it is RNA, and not contaminating DNA, that hybridized to the immobilized DNA. These data indicate that the *Ecotdk* segment in SVGT1-*Ecotdk* is transcribed and that the RNA transcripts are sufficiently stable to be detected by a four-hour labeling.

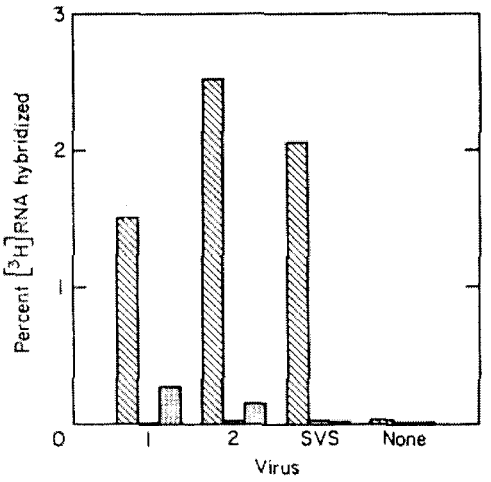


FIG. 5. Hybridization of ^{3}H-labeled RNAs obtained from CV-1 cells infected with SV40 and two SVGT1-*Ecotdk* recombinants.

Approximately 10^6 cts/min of each RNA were hybridized, separately, to nitrocellulose discs containing either SV40 (hatched), pTK1 (stippled), or no (open) DNA. The RNAs were from cells infected with: SVGT1-*Ecotdk* clone 6A plus *ts*A58 (1); SVGT1-*Ecotdk* clone 50A plus *ts*A58 (2); SV40 virus (SVS), and no virus. The amount of RNA hybridizing to each of the immobilized DNAs is expressed as a percentage of the total RNA.

The amount of *Ecotdk*-specific RNA made in the infected cells is about one-tenth as much as the SV40 RNA. The ratio of the *ts*A58 to SVGT1-*Ecotdk* genomes in comparable cultures, 48 hours after infection, was between 1 and 4 (probably reflecting the ratio of the two viruses in the innoculum) suggesting that there is about 20 to 50% as much RNA homologous to the *Ecotdk* DNA segment as to the SV40 DNA segment it replaced. This extent of transcription is significant as there is about as much *tdk*-specific RNA in the infected cells (0·1 to 0·3%) as there is of abundant mRNAs in certain eukaryotic cells.

To examine the size and composition of the RNAs containing *Ecotdk* nucleotide sequences, RNA (50 to 100 μg) isolated 40 to 48 hours after infection with SV40 or SVGT1-*Ecotdk* (plus *ts*A58) was denatured with glyoxal, electrophoresed in a 1·5% agarose gel, stained with ethidium bromide and then transferred to a diazobenzyl-oxymethyl paper which covalently binds the RNA (Alwine *et al.*, 1977). The RNA imprints were incubated with either ^{32}P-labeled, nick-translated pTK1 or SV40 DNAs, and then autoradiograms were made to locate the positions of RNA sequences homologous to the two DNAs (Fig. 6). Figure 6(a) shows the ethidium bromide-stained RNA obtained from infections with four independent SVGT1-*Ecotdk* isolates, two having the cloned segment with one orientation and the other two, the opposite orientation. The 18 S and 28 S ribosomal RNA species are clearly visible in all the samples. Figure 6(c) shows the imprints of the same four electrophoresed RNA samples after hybridization with the ^{32}P-labeled SV40 DNA probe; the late 16 S and 19 S mRNAs are apparent in each of the samples. Hybridization of similar RNA imprints with ^{32}P-labeled pTK1 DNA (Fig. 6(b)) produces only a diffuse labeling pattern over most of the tracks (the bare areas correspond to the positions of the ribosomal RNAs which reduce or prevent the transfer of RNAs having the same electrophoretic mobility to the diazobenzyloxymethyl paper). This result indicates

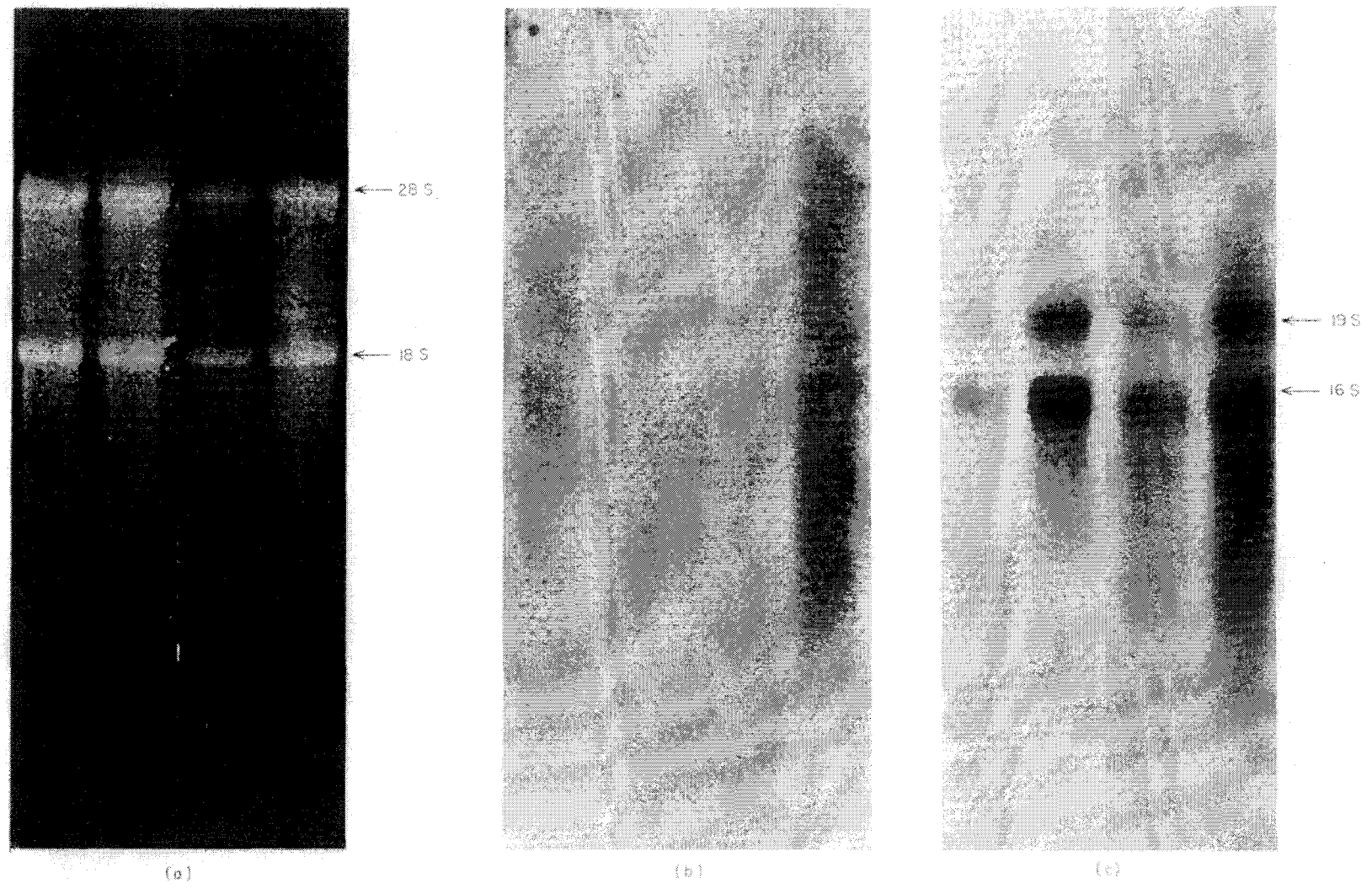


FIG. 6. Electrophoretic characterization of RNAs obtained from CV-1 cells infected with SVGT1-*Ecotdk* recombinants.
(a) The electrophoresed RNA was stained with ethidium bromide and the various RNA species were visualized with u.v. light. Tracks 1 to 4 contain RNAs extracted from cells infected with SVGT1-*Ecotdk* virus clones 22B, 27B, 47B and 50B, respectively. (b) The autoradiogram of RNAs complementary to ^{32}P-labeled pTK1 DNA. (c) The autoradiogram of RNA complementary to ^{32}P-labeled SV40 DNA.

that SVGT1-*Ecotdk* infected cell RNA does not contain discrete sized RNAs homologous to the *Ecotdk* DNA segment; rather, the *tdk* RNA sequences occur in a heterogeneous population of molecules ranging in size from about 1 to 8 kb.

Since most of the heterogeneous *Ecotdk*-specific RNA is larger than the 1·85 kb cloned *Ecotdk* segment, it is likely that these RNA molecules contain covalently joined SV40 and *Ecotdk* nucleotide sequences. To test this, ^{3}H-labeled RNA from cells infected with two of the recombinant viruses (plus *ts*A58) was annealed to nitrocellulose discs containing pTK1 DNA and, after extensive washing, the hybridized RNA was eluted (see Materials and Methods), precipitated with ethanol and reannealed to nitrocellulose discs containing either pTK1 or SV40 DNA. In two experiments, with different recombinant viruses, there was 30 to 40% as much hybridization of the pTK1 DNA-selected ^{3}H-labeled RNA to SV40 as to pTK1 DNA. Thus, at least 30% of the RNA homologous to the *Ecotdk* DNA is linked to RNA homologous to SV40 DNA. This is only a minimum estimate because *Ecotdk* and SV40 RNA sequences may have become separated by degradations during the first hybridization reaction.

Since *Ecotdk* DNA is transcribed during infection with the SVGT1-*Ecotdk* recombinants, we assayed the infected cell extracts for *E. coli* thymidine kinase activity. CV-1 cells and a TK$^-$ variant of monkey cells, T22TK$^-$, were infected with various SVGT1-*Ecotdk* hybrids (plus *ts*A58) and 30 hours later extracts were prepared and assayed for thymidine kinase (Table 1). The basal level of thymidine kinase activity in uninfected CV-1 cell extracts is increased about tenfold after SV40 infection (see Postel & Levine, 1976; Kit *et al.*, 1966). But the level of thymidine kinase activity

TABLE 1

*Thymidine kinase activities in extracts of CV-1 and T22TK$^-$ cells infected with SVGT1-*Ecotdk *recombinants*

Infecting virus	dTMP (pmol)	
	CV-1	T22TK$^-$
None	20	2
SV40	200	2
Clone 6A	130	3
Clone 22A	190	—
Clone 28A	250	3
Clone 47B	160	—
Clone 50A	90	—

CV-1 or T22TK$^-$ cell cultures were infected with wild-type SV40 or with various clones of SVGT1-*Ecotdk* recombinant viruses (plus *ts*A58 helper) and extracts were prepared as described in Materials and Methods either 30 h after infection at 41°C (CV-1) or 48 h after infection at 37°C (T22TK$^-$). All extracts contained between 0·75 to 1·0 mg protein/ml extract (Lowry *et al.*, 1951) and in each instance 30-μl portions were assayed for 30 min as described in Materials and Methods. In the assays with T22TK$^-$ extracts, the [^{3}H]thymidine was passed over a 0·5 ml column of DEAE-cellulose prior to use to remove contaminants which interferred with the determination of these very low levels of activity. The specific activity of the [^{3}H]thymidine was 5×10^2 cts/min per pmol and 10^2 cts/min per pmol in the assays with CV-1 and T22TK$^-$ extracts, respectively.

was not significantly higher after infection with five different isolates of SVGT1-*Ecotdk* (both orientations of the *Ecotdk* segment are represented in this set) than after infection with SV40 virus alone. When T22TK$^-$ monkey cells, which lack cytoplasmic thymidine kinase activity, were infected with either SV40 or two representative SVGT1-*Ecotdk* isolates, there was no reproducible increase in the level of thymidine kinase over the very low level seen in extracts of uninfected controls (Table 1). Concerned that even in TK$^-$ cells, the very low thymidine kinase activity was obscuring some formation of *E. coli* thymidine kinase, we resorted to gel electrophoresis to separate monkey and *E. coli* thymidine kinase. The infected cell extracts were electrophoresed and the thymidine kinase activity was assayed in the acrylamide gel (see Materials and Methods and Fig. 7). Under these conditions the R_F values of the mammalian cytoplasmic and mitochondrial activities are 0·1 and 0·2, relative to bromophenol blue, respectively; the R_F value of *E. coli*'s thymidine kinase is 0·7 and therefore, is well resolved from the mammalian enzymes. The results show that there is no detectable *E. coli* thymidine kinase activity in the tracks containing any of the SVGT1-*Ecotdk* infected cell extracts, although as little as 1 μl of the *E. coli* TK$^+$ extract gave a strong response (Fig. 7, track 7). We conclude that although the *Ecotdk* gene in SVGT1 is transcribed during multiplication of the recombinant genome, that RNA is probably not translated into active thymidine kinase.

(c) *Construction, characterization and expression of SVGT1*-ScetyrG *recombinant virus genomes*

S. cerevisae's haploid genome contains eight genes coding for tRNA^{Tyr} (Olson *et al.*, 1977), each one contained in a different size fragment after digestion of its DNA with *Eco*RI endonuclease (Goodman *et al.*, 1977; Olson *et al.*, 1979). Each of the eight *Eco*RI restriction fragments has been cloned in *E. coli* using λgtl as the vector (Olson *et al.*, 1979). The *ScetyrG* and *ScetyrE* DNA segments, 1·25 kb and 5·2 kb in length, respectively, are homologous to one another only over the approximately 120 bases coding for the tRNA^{Tyr} precursor; the nucleotide sequences surrounding the gene in *ScetyrG* and *ScetyrE* are different (Phillippsen, Cameron & Davis, unpublished results).

(i) *Cloning SVGT1*-ScetyrG *recombinants in monkey cells*

DNA segments greater than 2·2 kb cannot be accommodated in SVGT1: hence the 1·25 kb *ScetyrG* DNA segment was chosen for cloning with this vector. The *ScetyrG* DNA segment was excised from λgtl-*ScetyrG* with *Eco*RI endonuclease, purified by electrophoresis in agarose gel and its ends were "tailed" with poly(dT) (Goff & Berg, 1976). The modified *ScetyrG* fragment was annealed with SVGT1 DNA containing poly(dA) termini, *ts*A58 DNA was added and CV-1P cells were transfected with the mixture. Approximately 4000 plaques/μg of annealed DNA were produced and about 25% of these hybridized to ^{32}P-labeled *ScetyrG* DNA in the *in situ* plaque hybridization test.

DNA was prepared from infections of CV-1 cells with 60 plaques picked at random; 14 of these DNAs hybridized strongly to the ^{32}P-labeled λgtl-*ScetyrG* DNA probe. Virus in the plaques giving rise to these 14 DNA preparations were purified once by plaque isolation with added *ts*A58; then virus stocks and DNA were prepared from two isolates of each of the 14 original plaques. Seventeen of these 28 DNA preparations hybridized to the *tyrG* DNA probe.

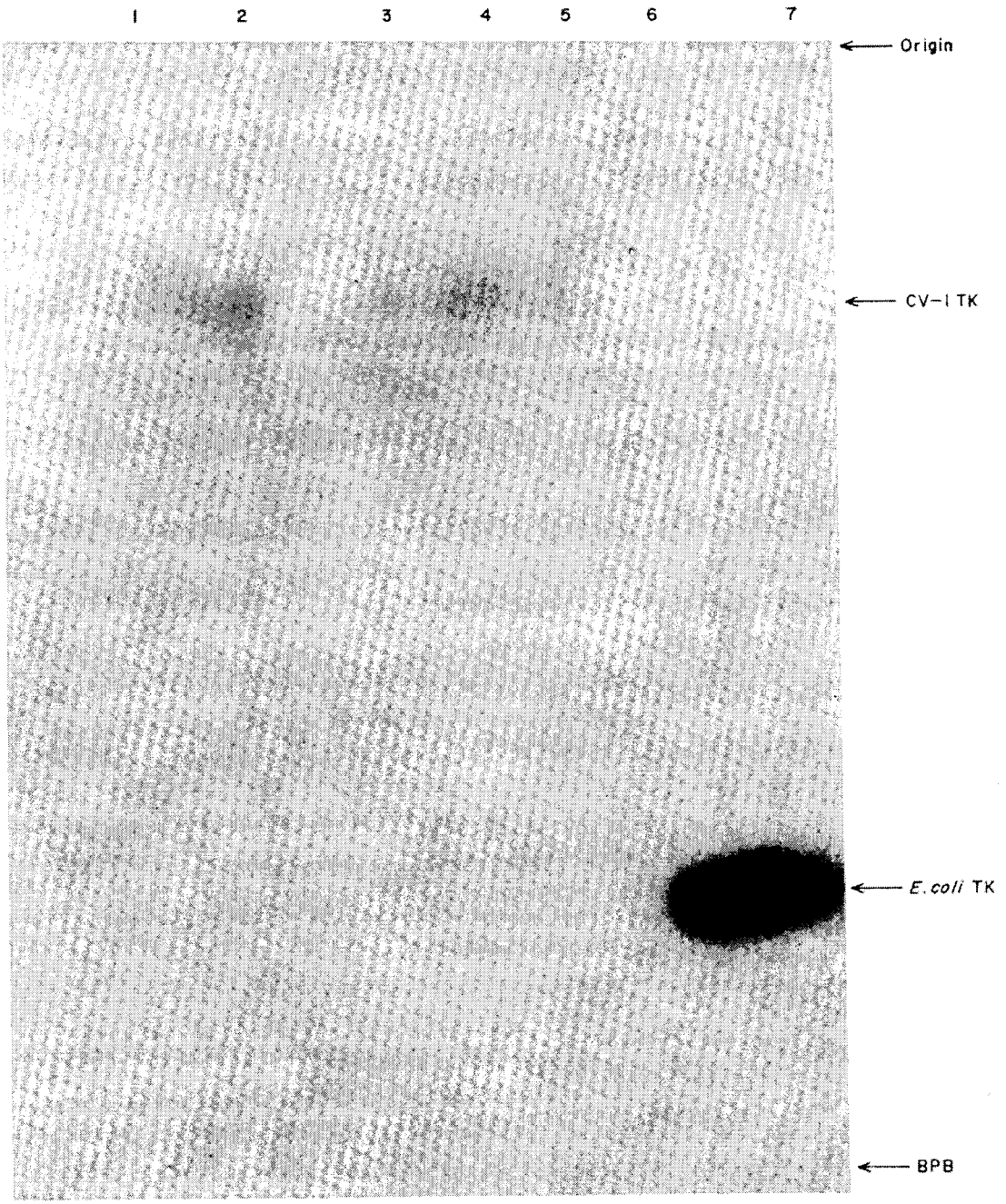


FIG. 7. Analysis of thymidine kinase in polyacrylamide gel after electrophoresis of extracts of CV-1 cells infected with various SVGT1-*Ecotdk* recombinants.

Extracts (30 μl) were loaded onto a 7% acrylamide gel and electrophoresed as described in Materials and Methods. The gel was soaked in a solution containing the assay ingredients, incubated and the ^{14}C-labeled dTMP formed precipitated *in situ*. The position of the dTMP was determined by autoradiography of the gels. The extracts were from cells infected with SVGT1-*Ecotdk* clones 50A, 44A, 27B, and 22A (plus their *ts*A58 helper virus) (tracks 1 to 4) or SV40 virus (track 5); uninfected cell extract was in track 6. Track 7 contained 1 μl of wild-type *E. coli* extract. The positions of the bromophenol blue marker (BPB), the CV-1 cytoplasmic and *E. coli* thymidine kinases are labeled.

(ii) *The structure of SVGT1-*ScetyrG *recombinant DNAs*

As with the SVGT1-*Ecotdk* recombinants, electrophoresis of restriction fragments in agarose gel was used to confirm the structure, location and orientation of the *ScetyrG* DNA segment in the recombinant genomes. Since the SVGT1 portion of the recombinant has only one *Taq*I restriction site and the *tyrG* segment contains five, digestion of the hybrids with *Taq*I endonuclease should produce four small fragments wholly from within the *tyrG* region (the arrangement of these four small fragments in the *tyrG* segment is not known); in addition, two large fragments, each containing part of the *tyrG* region, part of the SVGT1 vector, and a poly(dA:dT) join sequence should be formed (see Fig. 8(a)). Here too, the size of the large fragments is diagnostic of the orientation of the *tyrG* segment.

Of the 17 SVGT1-*ScetyrG* DNAs, six gave *Taq*I restriction digests consistent with the presence of a single recombinant plus helper DNA (Fig. 8(b)). Each of these generated full-length linear DNA from the *ts*A58 helper DNA and two additional large fragments from the recombinant (Fig. 8(b)). The sizes of these two fragments fell into two classes: they were either about 2·25 and 1·45 kb or 2·75 and 0·95 kb in length. These are the fragment sizes that would be obtained from recombinants having the *tyrG* segment inserted into the SVGT1 vector in either of the two possible orientations (Fig. 8(a)). Amongst the six uncomplicated clones (of which only three are independent), five contain the *tyrG* segment in one orientation and one has the *tyrG* segment in the opposite orientation. Here too, the small variations in the sizes of the join fragments (always less than 100 base-pairs) probably result from variations in the length of their poly(dA:dT) segments.

The four small fragments produced by *Taq*I endonuclease digestion of SVGT1-*ScetyrG* DNA derive from within the *tyrG* segment (Fig. 8). In each of the six isolates the four fragments had the same electrophoretic mobility as those produced by the original *tyrG* segment cloned in λgtl. We conclude that these recombinants contain faithful copies of the *ScetyrG* DNA. Since there are no comparable small fragments (40 to 200 base-pairs) produced from the vector or helper DNA and the resolution of the DNA bands under these electrophoretic conditions is high. a change of as few as five to ten base-pairs in these fragments could have been detected.

DNA from each of the remaining 11 cloned SVGT1-*ScetyrG* isolates was heterogeneous, since many additional fragments, some in less than molar yields, were found by gel electrophoresis of their *Taq*I endonuclease digests. Also, the internal *Taq*I restriction fragments from the *ScetyrG* segment were absent in two of these clones. Deletions and other rearrangements probably occurred during the isolation of these recombinant DNAs. But it is not clear whether the rearrangements occurred during the initial recombination event or during the subsequent propagation of the recombinant genomes.

(iii) *Expression of the* ScetyrG *DNA sequence*

From the orientation of the tRNA gene relative to the retriction map of the *tyrG* fragment (Olson *et al.*, 1979), we ascertain that the coding strand for $tRNA^{Tyr}$ is attached to the late strand of SV40 in five of the recombinants; therefore, transcription of SV40's L-strand (the predominant transcription late in SV40 infection) would also transcribe the template strand for the tRNA. The other SVGT1-*ScetyrG* recombinant has the template strand for $tRNA^{Tyr}$ joined to the early strand of the genome;

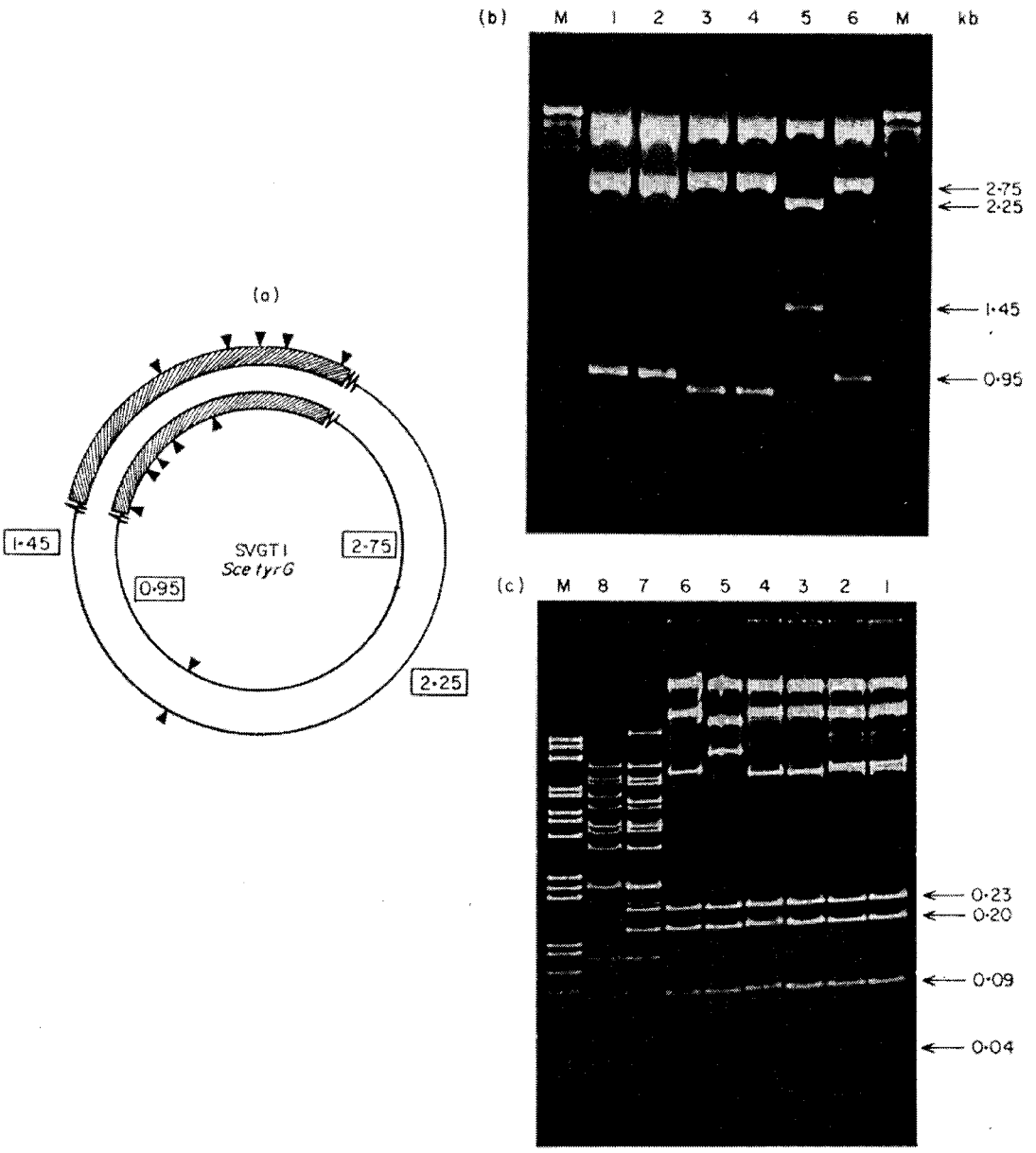


FIG. 8. *Taq*I endonuclease cleavage products of SVGT1-*ScetyrG* recombinant DNAs.

Various SVGT1-*ScetyrG* recombinant DNAs (plus *ts*A58 DNA) were cleaved with *Taq*I endonuclease and the resulting fragments separated by electrophoresis on agarose or polyacrylamide gels.

(a) *Taq*I endonuclease cleavage sites of SVGT1-*ScetyrG* showing both orientations of the *ScetyrG* segment.

(b) Analysis of the larger DNA fragments by electrophoresis on agarose gel. Tracks 1 to 6 correspond, respectively, to recombinant clones 61A, 61B, 64A, 64B, 97B, and 104B. The outer tracks (M) contain *Eco*R1 plus *Hin*dIII endonuclease-cleaved phage λ DNA as molecular length standards. In tracks 1 to 4 and 6 the recombinant DNAs yield the 2·75 and 0·95 fragments, indicative of the orientation shown in the inner circle of (a). The recombinant DNA in track 5 produces the 2·25 and 1·45 kb fragments, consistent with the orientation in the outer ring of (a).

(c) Analysis of the small DNA fragments by electrophoresis on polyacrylamide gel. Tracks 1 to 6 are as in (b); tracks 7 and 8 contain plasmid p*tyrG* and pMB9 DNAs, respectively, after cleavage with *Taq*I endonuclease. Track M has marker fragments produced by digestion of PM2 DNA with *Hae*III endonuclease.

consequently, transcription of the L-strand into the *tyrG* region of the recombinant DNA would yield anti-sense RNA, i.e. RNA complementary to the tRNA.

To explore whether the *tyrG* DNA segment is transcribed during infection of CV-1 cells with SVGT1-*ScetyrG* and, specifically, if a mature $tRNA^{Tyr}$ is produced, mock-infected, SV40 and SVGT1-*ScetyrG* (plus *ts*A58) cells were labeled with [^{3}H]uridine (4 h) 48 hours after infection and the total cellular RNA was isolated. The RNA from the uninfected and SV40-infected cells, did not hybridize to p*tyrG* DNA (Fig. 9); apparently, the nucleotide sequence of monkey $tRNA^{Tyr}$ is not homologous to the

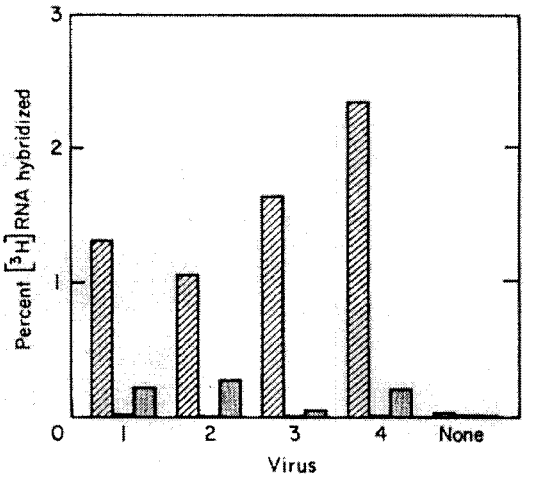


FIG. 9. Hybridization of ^{3}H-labeled RNAs obtained from CV-1 cells infected with SVGT1-*ScetyrG* recombinant viruses.

Approximately 10^6 cts/min of each RNA was hybridized separately to nitrocellulose discs containing either SV40 DNA (hatched), p*tyr*G DNA (stippled), or no DNA (open). The RNAs (sets 1 to 4) were from cells infected with SVGT1-*ScetyrG* clones 61A, 64A, 97B, and 104B, respectively, plus *ts*A58. The amount of RNA hybridizing to each of the immobilized DNAs is expressed as a percentage of the total RNA.

sequence in *S. cerevisae*'s $tRNA^{Tyr}$. However, ^{3}H-labeled RNA obtained from cells infected with the SVGT1-*ScetyrG* recombinants did hybridize to *ScetyrG* DNA (Fig. 9). As was found with SVGT1-*Ecotdk* infected cells, about 10% as much RNA was transcribed from the inserted DNA segment, as from the SV40 late region.

To determine if there were tRNA-sized RNAs made during infection with the SVGT1-*ScetyrG*, the RNA from infected cells was denatured with glyoxal, electrophoresed on 1·5% agarose gels, transferred to diazobenzyloxymethyl paper and annealed with ^{32}P-labeled p*tyrG* or SV40 DNA (Fig. 10). In addition to the diffuse labeling over the length of the strips, presumably because of the presence of heterogeneous sized RNA molecules (1 to 8 kb) homologous to the *ScetyrG* DNA, there is a discrete band with a mobility corresponding to 4 S RNA (Fig. 10(a), tracks 2 and 3). A comparable RNA preparation obtained from cells infected with SVGT1-*Ecotdk* did not hybridize with the *tyrG* DNA (Fig. 10(a), track 1). Annealing the imprints of the electrophoresed RNA with ^{32}P-labeled SV40 DNA established that each of the RNA preparations contained the expected 19 S and 16 S SV40 RNA species (Fig. 10(b), tracks 1 to 3). The same result, namely the formation of 4 S RNA homologous to nucleotide sequences in *ScetyrG*, was obtained in cells infected with SVGT1-*ScetyrG* recombinants whose *tyrG* segment occurred in either of the two possible orientations.

To establish that the 4 S species that hybridizes with *ScetyrG* DNA is homologous

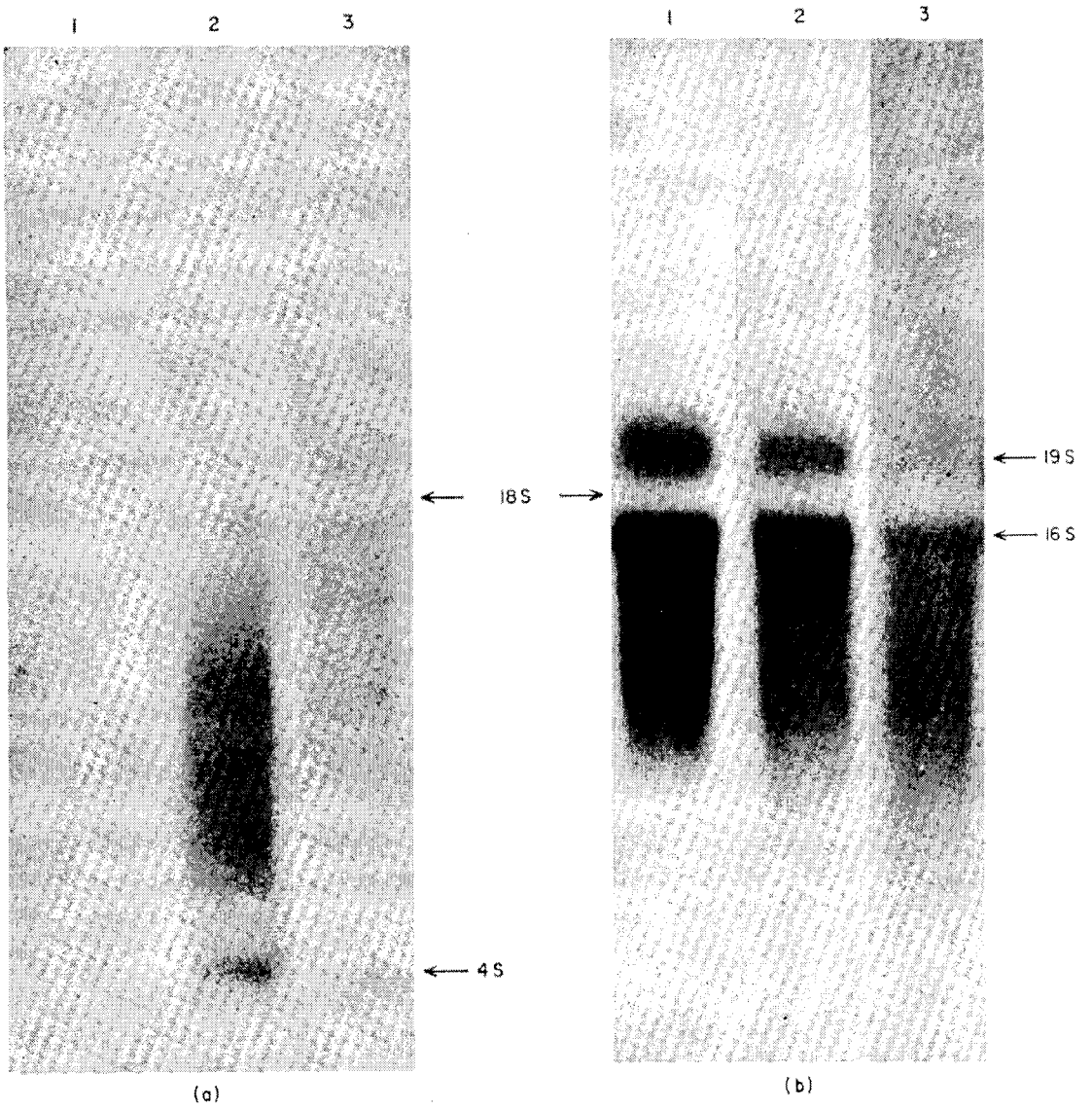


FIG. 10. Electrophoretic separation and hybridization of RNAs produced after infection with SVGT1-*ScetyrG* recombinant viruses.

RNAs obtained from CV-1 cells after infection with the indicated virus stocks were denatured with glyoxal, electrophoresed, and transferred to diazobenzyloxymethyl paper; then, the RNA imprints were annealed with ^{32}P-labeled p*tyrG* DNA (a) and subsequently with ^{32}P-labeled SV40 DNA (b). The RNA in track 1 is from an infection with SVGT-*EcotdK* clone 6A and in tracks 2 and 3 from infections with SVGT1-*ScetyrG* clones 61A and 97B.

to the nucleotide sequence coding for tRNA^{Tyr} and not its flanking sequences, the RNA was hybridized with ^{32}P-labeled *ScetyrE* DNA; this probe is specific for the tRNA^{Tyr} nucleotide sequence *per se*, since the *tyrE* and *tyrG* DNA segments are homologous only in their tRNA coding sequences and differ in their flanking sequences (Phillippsen, Cameron & Davis, unpublished results). RNA from SV40 and SVGT1-*ScetyrG*-infected cells was isolated, processed as described above and electrophoresed

in 2·5% agarose for 3·5 hours. Following transfer of the RNA to the diazobenzyloxymethyl paper, the samples were annealed with ^{32}P-labeled *tyrE* DNA and autoradiographed as usual (Fig. 11). RNA homologous to *tyrE* is clearly evident in the sample obtained from cells infected with the recombinant but not from cells infected with SV40. The most intensely labeled band has the same electrophoretic mobility as the 4 S or tRNA species but further studies are required to establish whether the 4 S RNA species made during infection with SVGT1-*SectyrG* recombinants is, indeed, fully modified, functional tRNA^{Tyr}.

With labeled *tyrE* DNA as the probe, there was also diffuse labelling of large RNA, though less than was found in the hybridization with *tyrG* DNA. This suggests that the entire cloned *tyrG* segment is probably transcribed. Moreover, at least 20% of the RNA that anneals to *tyrG* DNA is also homologous to SV40 DNA, a finding that suggests that some *tyrG*- and SV40-specific RNA nucleotide sequences are covalently joined, probably in a common RNA transcript.

4. Discussion

Our long term objective is to construct, *in vitro*, specialized transducing viruses that can multiply and express the transduced genetic information in mammalian cells. In the experiments reported here two genes, one coding for thymidine kinase in *E. coli* (*Ecotdk*) and the other specifying the structure of a *S. cerevisae* tRNA^{Tyr} (*ScetyrG*) have been transduced into cultured CV-1 monkey cells using a sub-genomic segment of SV40 DNA (SVGT1) as the vector.

SVGT1 is a 3·05 kb DNA segment obtained by cleavage of SV40 DNA with *Mnol* and *Bam*H1 endonucleases followed by electrophoresis in agarose. The vector DNA contains the entire early region and the origin of DNA replication, but lacks virtually all of the coding information for three virion proteins. Consequently, although SVGT1's DNA can replicate in permissive cells (e.g. CV-1 monkey cells) and transform non-permissive cells (e.g. rodent cells), it cannot produce mature virions. However, SVGT1 and its recombinants can be propagated by complementation with viral genomes that contain a complete, functional late region. In our protocol an early mutant of SV40, *ts*A58, provides the late functions; *ts*A58's replication defect is, in turn, complemented by SVGT1 (Goff & Berg, 1976).

It is tedious, time consuming and costly to screen and recover recombinant genomes if the segments to be cloned in SVGT1 represent less than 10^{-2} of the DNA population. To circumvent this difficulty we have resorted to molecular cloning with bacteriophage and plasmid vectors in *E. coli*, a relatively rapid and inexpensive way to purify relevant DNA segments for recombination with SVGT1. By cloning in *E. coli*, the desired DNA segment can be trimmed to the proper size, purified from extraneous DNA sequences and prepared for *in vitro* recombination under conditions which assure efficient formation, screening and recovery of the particular recombinants. For example, using ϕ80 phage and pMB9 as vectors for molecular cloning in *E. coli*, the 1·85 kb *Ecotdk* DNA segment inserted into SVGT1-*Ecotdk* was purified nearly 2000-fold from approximately 4000 kb of *E. coli* DNA. The *tdk* gene was particularly suitable for this approach because its presence could be selected for in TK$^-$ host cells. In the absence of such a selection for function, the desired segment could be detected by other means, e.g. by hybridization with appropriate labeled RNA or DNA probes (Grunstein & Hogness, 1975; Benton & Davis, 1977) or by relying on the

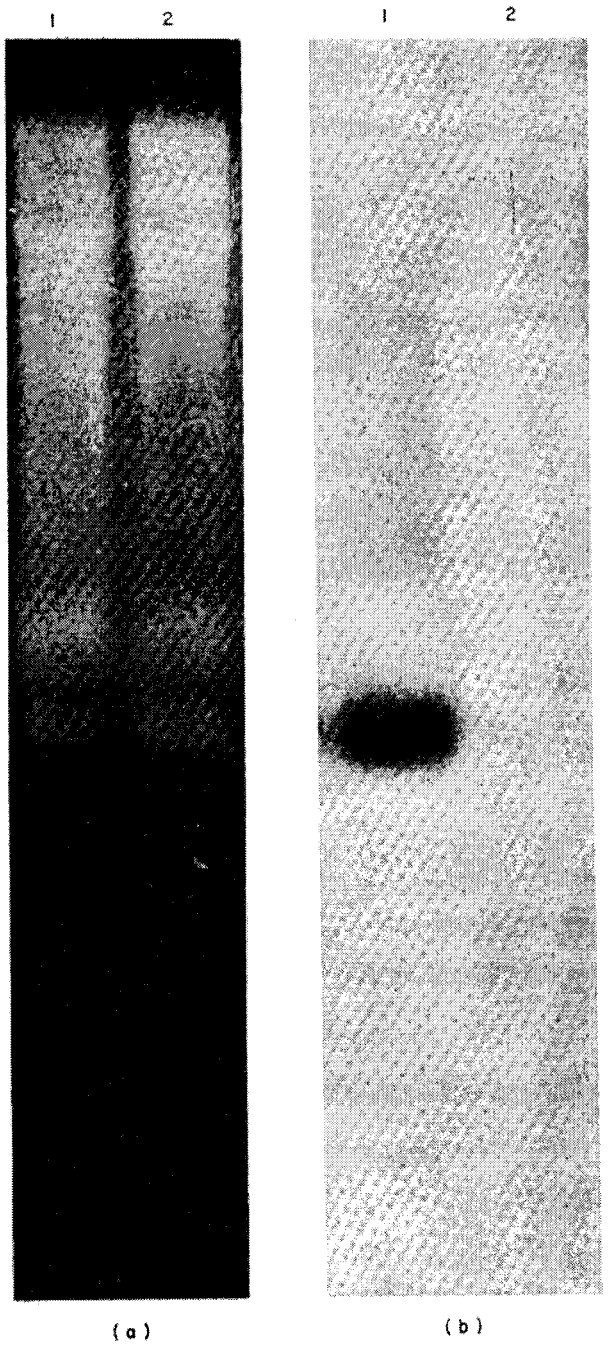


FIG. 11. Electrophoresis and hybridization of RNA made after infection with a SVGT1-*ScetyrG* recombinant virus.

(a) RNA from CV-1 cells infected with SVGT1-*ScetyrG* clone 61A (track 1) or SV40 (track 2) was denatured and electrophoresed on 2·5% agarose. The gels were stained with ethidium bromide: the 18 and 28 S ribosomal RNAs are at the top and the 5 and 4 S RNAs in the middle of the gel. (b) The different RNA species were transferred to diazobenzyloxymethyl paper, hybridized with ^{32}P-labeled pMB9-*ScetyrE* DNA and the paper was autoradiographed.

antigenicity of the gene products (Skalka & Shapiro, 1976). In constructing SVGT1-*ScetyrG*, the purified *ScetyrG* segment was already available as the λgtl-*ScetyrG* recombinant (Olson *et al.*, 1979). *E. coli*'s gene for guanine-phosphoribosyl-transferase (*Ecogpt*) has also been recombined with SVGT1 (Mulligan & Berg, unpublished results) using DNA from a *λpgpt* transducing phage (Harriman & Mulligan, unpublished results) as the starting material; similarly, the plasmid cDm500 (Lifton *et al.*, 1977) provided purified *Drosophila melanogaster* histone genes for recombination with SVGT vectors (Hofstetter & Berg, unpublished results). Clearly, the "banks" and "libraries" of cloned pro- and eukaryote genes that are accumulating, provide a rich source of purified DNA segments for insertion into appropriate vectors and transduction into animal cells.

To interpret the expression of the transduced DNA segments we need to consider, briefly, our present understanding of SV40's late transcription. SV40's capsid polypeptides VP1, VP2 and VP3 are coded by mRNAs transcribed from the DNA between map co-ordinates 0·67 to 0·17. (Fig. 12 shows the RNAs as open bars and the coding segment for each polypeptide as a filled bar.) The late region is transcribed in the clockwise direction (Khoury *et al.*, 1973,1975) ultimately yielding mRNAs with polyadenylated 3′ ends at about map position 0·175. The 5′-termini of the late mRNAs are heterogeneous and map between co-ordinates 0·67 to 0·73 (Lai *et al.*, 1978); it is unclear whether the different 5′-termini arise by processing events following initiation of transcription at a unique promoter site, or by initiation of transcription at many sites. A feature of SV40's late mRNAs is that their nucleotide sequences are not contiguous with respect to the DNA (see Fig. 12). The major mRNA coding for VP1, for example, contains a "leader" sequence of about 200 to 210 nucleotides homologous to map positions 0·725 to 0·76, joined to an RNA segment transcribed

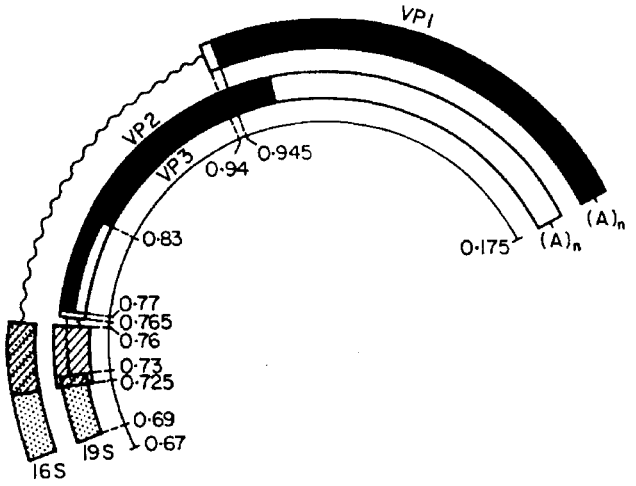


FIG. 12. Representation of the structure of SV40's late mRNAs.

The map of SV40 DNA from the origin of DNA replication (0·67) to the end of the late region (0·175) is shown on the inner circle. The regions coding for the structure of VP1, VP2, and VP3 are shown as shaded regions within bars that define the "bodies" of the 19 S and 16 S mRNAs.

The "leader" segments are shown as bars spanning map co-ordinates 0·69 to 0·76. The hatched and stippled regions of the 16 S mRNA leader, for example, are intended as symbolic, rather than literal, representations of more than one type of leader segment; one leader segment spans the region from 0·725 to 0·76 (hatched) and another the region from 0·69 to 0·76 (stippled). For the 19 S mRNA species one class of leaders spans map co-ordinates 0·69 to 0·73 (stippled) and is joined to the body at 0·76; another class of leaders extends from 0·725 to 0·76 (hatched).

from map positions 0·935 to 0·175: the intervening RNA sequence from map position 0·76 to 0·935 is missing from the VP1 mRNA (Aloni *et al.*, 1977).

Mutants with deletions that extend through the region between map co-ordinates 0·75 and 0·79 fail to produce any of the late proteins or their cytoplasmic mRNAs (R. Mulligan, L. Villarreal, R. White & P. Berg, unpublished results). Since nuclear RNA homologous to the late region is synthesized, the absence of mature cytoplasmic mRNAs probably results from a failure to effect proper "splicing". Thus, the region from 0·75 to 0·79 seems to have a crucial role in the processing of late nuclear RNA to functional mRNAs.

The recombinant genome, SVGT1-*Ecotdk*, lacks the nucleotide sequences distal to 0·73. Consequently, although transcription of the *Ecotdk* DNA segment may be initiated at one or more sites proximal to map position 0·73, the transcripts seem not to be processed into stable cytoplasmic RNAs. The heterogeneous sized RNA species which contain covalently joined SV40 and *Ecotdk* nucleotide sequences, may be molecules at various stages in the turnover of the initial transcript. A similar pattern of λ-specific RNA accumulation was found (Howard & Berg, unpublished results) during a re-examination of the RNAs produced during infection with a previously reported recombinant genome, SVGT1-λ (Goff & Berg, 1976).

The situation with the SVGT1-*ScetyrG* recombinant is different and intriguing. Here, too, transcription probably originates in the SV40 DNA and continues into the *ScetyrG* segment as evidenced by the appearance of large, heterogeneous sized RNA homologous to the $tRNA^{Tyr}$ and flanking regions of the *tyrG* segment. But in this instance a discrete RNA species is also produced, an RNA whose electrophoretic mobility resembles a mature 4 S tRNA species and which is homologous, specifically, to the region in the *ScetyrG* segment that codes for $tRNA^{Tyr}$. Further studies are in progress to determine if this RNA is spliced (Goodman *et al.*, 1977), modified and biologically functional yeast $tRNA^{Tyr}$. Of particular interest is whether this yeast RNA species is generated by processing of an RNA transcript that had been initiated on SV40 DNA, or whether its transcription begins within the cloned segment, perhaps at the *ScetyrG* promoter itself. Consistent with the latter possibility is our observation that the same tRNA-like species is produced after infections with SVGT1-*ScetyrG* recombinants having the *tyrG* segment inserted in both orientations. Quite possibly the monkey cell's RNA polymerase III can initiate transcription of the yeast tRNA genes; RNA polymerase III from several eukaryotes can transcribe 5 S and 4 S genes from unrelated species (Parker *et al.*, 1977; Schmidt *et al.*, 1978).

Summarizing, our findings suggest that transcription of DNA segments inserted into SVGT1 occurs, but the failure to process these transcripts properly prevents the accumulation of stable large mRNAs. The recognition of post-transcriptional RNA splicing and its importance in producing stable cytoplasmic mRNAs places new constraints on the structure of the SV40 DNA vector. Quite clearly, SVGT1 is not a useful vector for transducing DNA segments which lack a transcriptional promoter or sites which would permit post-transcriptional processing of the joint SV40-"foreign" DNA primary transcript. For DNA segments lacking such "signals", but containing an intact protein-coding sequence, an alternative strategy is possible. SV40 DNA vectors have been constructed which can accommodate foreign genes within one of SV40's genes, thereby permitting the coding sequences of the transduced genes to be included within the "body" of the corresponding SV40 mRNAs (Mulligan, Howard & Berg, and Hofstetter & Berg, unpublished results).

This work was supported by United States Public Health Service research grant GM 13235 from the National Institute of General Medical Sciences and by grant VC23 from the American Cancer Society. One of us (S. P. G.) was a Smith, Kline, and French predoctoral scholar, as well as a United States Public Health Service Trainee.

REFERENCES

Aloni, Y., Dhar, R., Laub, O., Horowitz, M. & Khoury, G. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 3686–3690.

Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5350–5354.

Bachmann, B. J., Low, K. B. & Taylor, A. L. (1976). *Bacteriol. Rev.* **40**, 116–167.

Benton, D. W. & Davis, R. W. (1977). *Science*, **196**, 180–182.

Bolivar, F., Rodriguez, R. L., Betlach, M. C. & Boyer, H. W. (1977). *Gene*, **2**, 75–93.

Cameron, J. R. (1977). Ph.D. thesis, Stanford University, Stanford.

Davis, R. W., Simon, M. & Davidson, N. (1971). In *Methods in Enzymology* (Grossman, L. & Moldave, K., eds), vol. 21D, pp. 413–428, Academic Press, New York.

Davoli, D., Nussbaum, A. L. & Fareed, G. C. (1976). *J. Virol.* **19**, 1100–1106.

Ganem, D., Nussbaum, A. L., Davoli, D. & Fareed, G. C. (1976). *Cell*, **7**, 349–359.

Goff, S. P. & Berg, P. (1976). *Cell*, **9**, 695–705.

Goff, S. P. & Berg, P. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 1763–1767.

Goodman, H. M., Olson, M. V. & Hall, B. D. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5453–5457.

Grunstein, M. G. & Hogness, D. S. (1975). *Proc. Nat. Acad. Sci., U.S.A.* **72**, 3961–3965.

Hamer, D. (1977). In *Recombinant Molecules: Impact on Science and Society* (Beers, R. G., Jr & Bassett, E. G., eds), pp. 317–335, Raven Press, New York.

Hirt, B. (1967). *J. Mol. Biol.* **26**, 365–369.

Katz, L., Kingsbury, D. T. & Helinski, D. R. (1973). *J. Bacteriol.* **114**, 557–591.

Kit, S., Dubbs, D. R., Frearson, P. M. & Melnick, J. L. (1966). *Virology*, **29**, 69–83.

Khoury, G., Martin, M. A., Lee, T. N. H., Danaa, K. J. & Nathans, D. (1973). *J. Mol. Biol.* **78**, 377–389.

Khoury, G., Howley, P., Nathans, D. & Martin, M. (1975). *J. Virol.* **15**, 433–437.

Lai, C.-J., Dhar, R. & Khoury, G. (1978). *Cell*, **14**, 971–982.

Lifton, R. P., Goldberg, M. L., Karp, R. W. & Hogness, D. S. (1977). *Cold Spring Harbor Symp. Quant. Biol.* **42**, 1047–1052.

Lowry, O. H., Rosebrough, N. J., Farr, A. C. & Randall, R. J. (1951). *J. Biol. Chem.* **193**, 265–275.

McMaster, G. K. & Carmichael, G. G. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 4835–4838.

Mertz, J. E. & Berg, P. (1974). *Virology*, **62**, 112–124.

Nussbaum, A. L., Davoli, D., Ganem, D. & Fareed, G. C. (1976). *Proc. Nat. Acad. Sci., U.S.A.* **73**, 1068–1072.

Okazaki, R. & Kornberg, A. (1964*a*). *J. Biol. Chem.* **239**, 269–274.

Okazaki, R. & Kornberg, A. (1964*b*). *J. Biol. Chem.* **239**, 275–284.

Olson, M. V., Montgomery, D. L., Hopper, A. K., Page, G. S., Horodyski, F. & Hall, B. D. (1977). *Nature (London)*, **267**, 639–641.

Olson, M. V., Hall, B. D., Cameron, J. R. & Davis, R. W. (1979). *J. Mol. Biol.* **127**, 285–295.

Parker, C. S., Jaehning, J. A. & Roeder, R. G. (1977). *Cold Spring Harbor Symp. Quant. Biol.* **42**, 577–588.

Peacock, A. C. & Dingman, C. W. (1968). *Biochemistry*, **7**, 668–674.

Postel, E. H. & Levine, A. J. (1976). *Virology*, **73**, 206–215.

Radloff, R., Bauer, W. & Vinograd, J. (1967). *Proc. Nat. Acad. Sci., U.S.A.* **57**, 1514–1521.

Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977). *J. Mol. Biol.* **113**, 237–251.

Schmidt, O., Mao, J.-I., Silverman, S., Hovemann, B. & Söll, D. (1978). *Proc. Nat. Acad. Sci., U.S.A.* **75**, 4819–4823.

Skalka, A. & Shapiro, L. (1976). *Gene*, **1**, 65–79.

Tischfield, J. A., Bernhard, H. P. & Ruddle, F. H. (1973). *Anal. Biochem.* **53**, 545–554.

Villarreal, L. P. & Berg, P. (1977). *Science*, **196**, 183–185.

Wensink, P. C., Finnegan, D. J., Donelson, J. E. & Hogness, D. S. (1974). *Cell*, **3**, 315–325.